

Evaluating the Quality of Spatial Data for the Analysis of Climate Variability in the Coastal Region of Nigeria

Agbonaye A. I.^{1,*} and Izinyon O. C.²

^{1,2}Department of Civil Engineering, University of Benin, Benin City, Nigeria

Corresponding Author: *augustine.agbonaye@uniben.edu

<https://doi.org/10.36263/nijest.2021.01.0236>

ABSTRACT

The lack of truly reliable data for climate change analyses and prediction presents challenges in climate modeling. Needed data are required to be hydrologically/statistically reliable to be useful for hydrological, meteorological, climate change, and estimation studies. Thus, data quality and homogeneity screening are preliminary analyses. In this study, the homogeneity of the climatic data used for analyses of climate variability was conducted in the coastal region of Nigeria. Climatic Research Unit (CRU 0.5× 0.5) gridded monthly climatic data for sixty years (1956- 2016) for nine states of the coastal region of Nigeria obtained from internet sources were validated with the Nigerian Meteorological Agency (NiMet) data to assure adequacy for use. The data were tested for normality using the Shapiro-Wilk (S-W) test, D'Agostino-Pearson omnibus test, and skewness test. Four homogeneity test methods were applied to 257 locations in the nine states of the coastal region of Nigeria and they include Pettit's, Standard Normal Homogeneity Test (SNHT), Buishand's and Von Neumann Ratio (VNR) tests. The results of the validity analysis indicated that the CRU data are very reliable and thus justified their use for the further analysis carried out in the study. Also, the results obtained indicated that CRU climatic data series were normally distributed and parametric methods could be used in further analysis of the data. Rainfall data homogeneity was detected for Bayelsa, Delta, Edo, Lagos, Ogun, and Ondo states and inhomogeneity for Akwa Ibom, Cross Rivers, and Rivers States. Also, temperature data inhomogeneity was detected for all the states in the study area.

Keywords: Climatic data, Normality, Homogeneity, Climate change, CRU

1.0. Introduction

Reliable and accurate estimates of climate are not only crucial for the study of climate variability but are also important for water resource management, agriculture, weather, climate, and hydrological forecasting (Sarojini *et al.*, 2016). Unfortunately, there is the dearth of satisfactory climatic data in developing countries. The rain-gauge measurement is the traditional and oldest method for monitoring rainfall. However, because of the practical observational limitations, this measurement often suffers from numerous gaps in space and time, due to weather stations being limited in numbers and often unevenly distributed, resulting in missing data problem, a short period of observation, incomplete areal coverage, and deficiencies over most oceanic and sparsely populated areas (Kidd *et al.*, 2017), thus making its use in climate change diagnostic studies less reliable in initial data processing and calibration problems of subjecting non-continuous rainfall and temperature data into the Water Balance and TREND software. This may arise as this software often recognizes only continuous data of long duration over fifty (50) years. Unfortunately, available data from NIMET either had missing data or are not up to 50 years in some locations. Often, this basic requirement for use is not met and the situation is compounded by sparse data coverage challenge in the region due to the few weather stations per state most of which are located in the airports. Studies have shown that a major constraint in climatic change research identified in the past is the lack of long-term climatic data on a temporal basis (Nnamchi *et al.*, 2008). Mitchel and Jones (2005) recommended that a large proportion of such data needs can be met through providing a standard set of 'climate grids', in terms of monthly variations over a century-long time scale on a regular high-resolution (0.5°) latitude-longitude grid.

The extensive application of data the CRU TS datasets for studying climatic variability has been used by converting climatic datasets into formats that are more commonly used and therefore can be directly utilized in GIS. In long-term modeling of climatic variability, the application of the CRU dataset is well documented.

Data taken from the observation stations should be tested for reliability and homogeneity before their use in the research studies. Homogeneity testing is one of the most important analyses in climate-related studies as it underpins the reliability of any inferences (Yozgatligil and Yazici, 2015). The accuracy and reliability of the model result in the studies about climate change, classification, flood and drought modeling, water resources planning and modeling related to hydrology and meteorology vary according to the quality of the data used. Inhomogeneous observation records may occur in the stations making meteorological observations unreliable due to the method used, the conditions around the station and the reliability of the measurement tool, etc.

Homogeneous climate series may be defined as a series only influenced by the variations in climate. However, it is difficult to find reference stations with a high correlation and a homogeneous structure in wide regions. For this reason, the absolute method was used for the homogeneity test in our study owing to the high spatial variation of precipitation stations. The standard procedure is to apply four tests: the Pettit, Standard Normal Homogeneity (SNH), Buishand (BR), and Von Neumann (VNR) tests at a significance level of 0.05 (Agha *et al.*, 2017). The temperature series were tested using the annual mean temperature while the Precipitation series were tested using the annual mean rainfall. The use of derived annual variables avoided autocorrelation problems with testing daily series.

There are some differences between SNHT, BRT, and Pettitt test. SNHT test is known to find change point towards the beginning and the end of the series, whereas BRT and Pettitt tests are sensitive to find the changes in the middle of a series (Martínez *et al.*, 2009). These three tests are capable of detecting the year where a break occurs. Meanwhile, VNRT assumes the same null hypothesis as the previous three tests but for the alternate hypothesis, it assumes that the series is not randomly distributed. VNRT assesses the randomness of the series but does not give information about the year of the break. Homogeneity of consistency implies that all the collected hydrologic time series data belong to the same statistical population having a time-invariant mean. Therefore, the tests to check the homogeneity or consistency of the data series are based on evaluating the significance of changes in the mean value. To be accurate, the climate data used for long-term climate analyses, particularly climate change analyses, must be homogeneous. A homogeneous climate time series is defined as one where variations are caused only by variations in weather and climate. Unfortunately, most long-term climatological time series have been affected by several non-climatic factors that make these data unrepresentative of the actual climate variations occurring over time.

Climate data homogenization aims to adjust observations, if necessary so that the temporal variations in the adjusted data are caused only by climate processes.

2.0. Methodology

2.1. Description of study area

The study area is the coastal region of Nigeria. The area is geographically located between latitude 4°N – 8°N and longitude 3°E – 9°E. Figure 1 presents the map of Nigeria showing the coastal region station coordinates of representative cities in the coastal region of Nigeria and other states details are presented in Table 1. The Nigeria coastline which is about 853km long runs through nine states of Nigeria namely: Lagos, Ogun, Ondo, Edo, Delta, Bayelsa, Rivers, Akwa-Ibom and Cross River states, bordering the Atlantic Ocean. Nine stations, each from these states, were selected for representative coverage. Nigerian coastal zone experiences a tropical climate consisting of the rainy season (April to November) and dry season (December to March). High temperatures and humidity as well as marked wet and dry seasons characterize the Nigerian climate. The coastal areas have an annual rainfall ranging between 1, 500, and 4,000 mm. The coastline area is humid with a mean average temperature of 24 – 32°C (Kuruk, 2004). The coastal area is low lying with heights of not more than 3.0m above sea level and is generally covered by freshwater swamp, mangrove swamp, lagoon marshes, tidal channels, beach ridges, and sand bars (Nwilo and Badejo, 2006).

The station coordinates of representative cities in the coastal region of Nigeria and the map of the area are presented in Table 1 and Figure 1 respectively.

Table 1: Station coordinates of representative cities in the coastal region of Nigeria and other states details

S/N	Coastal state	Land Area (km ²)	Population density (Persons/km ²)	Selected city	Longitude	Latitude	Years of available MIMET data
1	Akwa-Ibom	8,421	466	Uyo	7.53°E	5.10°N	60
2	Bayelsa	21,100	81	Yenogoa	6.26°E	4.92°N	60
3	Cross Rivers	23,074	125	Calabar	8.20°E	4.57°N	60
4	Delta	17,011	241	Warri	5.75°E	5.5°N	60
5	Edo	15,650	206	Benin City	5.31°E	6.20°N	60
6	Lagos	3577 (787 in water)	2520	Ikeja	3.45°E	6.2°N	60
7	Ogun	16720	223	Abeokuta	3.35°E	7.16°N	60
8	Ondo	14769	233	Akure	5.5°E	7.15°N	60
9	Rivers	11,225	462	Port Harcourt	7.10°E	4.4°N	60

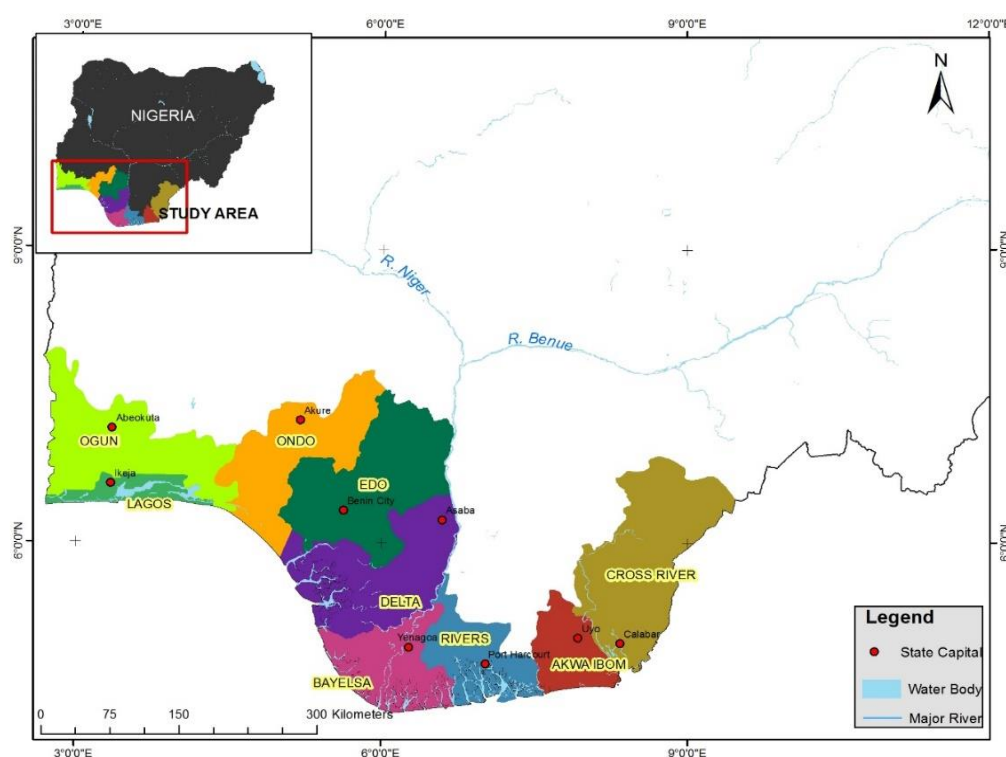


Figure 1: Map of Nigeria showing the coastal region

[Source: Adapted from Office of the Survey General of the Federation (OSGOF)]

2.2. Data collection/validation

Climatic data collected were mean monthly rainfall and temperature for a period of 60 years (1956-2017) for selected cities in the coastal region of Nigeria. Climate Research Unit (CRU 0.5× 0.5) gridded monthly climatic data for two climatic periods ((1956- 1986 and 1987-2016), obtained from the internet (<http://badc.nerc.ac.uk>), were sorted into annual rainfall series and validated with the Nigerian Meteorological Agency (NiMet) data obtained from Central Bank of Nigeria (CBN) website <https://www.cbn.gov.ng/documents/Statbulletin.asp>. The validation was carried out using the following goodness of fit statistics: Coefficient of Determination (R^2), Nash-Sutcliffe Efficiency (NSE) and Ratio of Standard Deviation of observations to Root Mean Squared (RSR). These data series from 257 locations in nine states of the region were tested for normality and homogeneity. The value of R^2 obtained is considered very good when its value is within the range of $0.75 < R^2 \leq 1$ (see Table 2). The lack of truly reliable data is a problem that complicates the analysis of climate trends, increasing the challenges of related relevant research, hence data validation was done.

Table 2: General validation rating used

R^2 (adj R_c) ²	RSR (-)	K tau	Validation rating	MAPE (%)	Validation rating	Cp	Capability
$0.75 \leq R^2 \leq 1$	$0 < \text{RSR} \leq 0.5$	$0.75 < K \leq 1$	Very Good	$0 < 10$	Highly accurate	$0 < 1$	Incapable
$0.65 \leq R^2 \leq 0.75$	$0.5 < \text{RSR} \leq 0.6$	$0.65 < K \leq 0.75$	Good	10-20	Good	$1 < \text{Cp} < 3$	capable
$0.5 \leq R^2 \leq 0.65$	$0.6 < \text{RSR} \leq 0.7$	$0.5 < K \leq 0.65$	Satisfactory	20-50	Reasonable	$3 < \text{Cp}$	Very capable
$R^2 \leq 0.5$	$\text{RSR} > 0.7$	$K \leq 0.5$	Unsatisfactory	> 50	Inaccurate		

Note that adjusted- R^2 can have a negative value, unlike R^2 , which is always between 0 and 1. We also mentioned that adjusted- R^2 is roughly equal to R^2 . Adjusted R-squared more than 0.75 indicates a very good value for showing the accuracy. The interpretations of Kendall's tau and Spearman's rank correlation coefficient are very similar and thus invariably lead to the same inferences.

2.3. Preliminary data analysis

2.3.1. Descriptive statistics

The descriptive statistics of annual rainfall time series which includes; minimum, maximum, range, mean, standard deviation, variance, standard error of the mean, kurtosis, and skewness with their standard error's computations were implemented by using XLSTAT software. Basic equations for descriptive statistics are presented in Table 3.

Table 3: Equations for descriptive statistics

Statistical text	Applicable formula	Remark
Mean	$\bar{R} = \frac{\sum(R)}{N}$	It is obtained by adding together all the variates, $\sum(R)$ and dividing by the total number of variates, N .
Standard Deviation	$\sigma = \sqrt{\frac{\sum(R - \bar{R})^2}{N}}$	It is the square root of the mean-squared deviation of individual observations from their mean Standard deviation.
Variance	$\sigma^2 = \frac{\sum(R - \bar{R})^2}{N}$	
Coefficient of Variation	Coefficient of variation, $C_v = \frac{\sigma}{\bar{R}}$	It is obtained by dividing the standard deviation of the data set by its mean.
Determination of coefficient of variation	$CV = \frac{SD \times 100}{\bar{X}}$	This is a measure of the variability of data
Skewness	$\alpha = \frac{1}{N} \sum(R - \bar{R})^3$	Where R is a variate, \bar{R} is the mean of the data set and N is the total number of variates.
Kurtosis	$K = n \frac{\sum_{i=1}^n (R_i - \bar{R})^4}{(\sum_{i=1}^n (R_i - \bar{R})^2)^2}$	Kurtosis is the degree of peakedness of distribution, usually taken relative to a normal distribution. A distribution having a relatively high peak is called leptokurtic, while a curve that is flat-topped is called platykurtic. The normal distribution which is not very peaked or very flat-topped is called isocratic.

2.3.2. Simple t -test

This test is data-driven and provided a known change point. It was used to check the null hypothesis of whether equal means in two different periods are different (Mu *et al.*, 2007). The simple t -test assumes that the data are normally distributed. The relationship is expressed as follows:

$$t = \frac{|\bar{x} - \bar{y}|}{S \sqrt{\frac{1}{n_1} + \frac{1}{n_2}}} \quad (1)$$

$$S = \sqrt{\frac{(n_1 - 1)s_1^2 + (n_2 - 1)s_2^2}{n - 2}} \quad (2)$$

where \bar{x} and \bar{y} in Equation 1 stand for the means of the observed and simulated data respectively, while n_1 and n_2 are the numbers of data in the observed and simulated respectively, and S is the standard deviation of sample (of the entire n_1 and n_2 observations, $n = n_1 + n_2$). We considered the t

statistic test statistic and the degrees of freedom to determine the p -value. The p -value is the probability that a t statistic having $n - 2$ (22) degrees of freedom is very much greater than 1.725. Since this is a two-tailed test, "very much greater" means greater than 1.725 or less than -1.725.

2.3.3. Test of normality

Normality test, which was employed to determine whether the dataset could be described by a normal distribution, was carried out using Shapiro Wilk Test (SWT), D'Agostino-Pearson Test, and Skewness Test. This enabled data screening, outlier identification, description, assumption checking, and characterizing differences among sub-populations.

The basic test of a hypothesis that guides in deciding on normality is stated as follows:

Null hypothesis H_0 : Data follow a normal distribution

Alternative hypothesis H_1 : Data do not follow a normal distribution

The Shapiro-Wilk

This test is one of the most popular tests for normality assumption diagnostics which has good properties of power and is based on correlation within given observations and associated normal scores. The Shapiro-Wilk test statistics derived by Shapiro and Wilk (1965) is as shown below:

$$W = \frac{\left(\sum a_i y_{(i)}\right)^2}{\sum (y - \bar{y})^2} \quad (3)$$

where $(i) y$ is the i^{th} order statistics and I_a is the i -th expected value of normalized order statistics. For independently and identically distributed observations, the values of I can be obtained from the table presented by Shapiro and Wilk (1965) for sample sizes up to 50. W can be expressed as a square of the correlation coefficient between I_a and $(i)y$. So W is the location and scale-invariant and is always less than or equal to 1. In the plot of $(i)y$ against I an exact straight line would lead to W very close to 1. So if W is significantly less than 1, the hypothesis of normality will be rejected. Although the Shapiro-Wilk W test is very popular, it depends on the availability of values of I_a , and for large sample cases their computation may be much more complicated.

D'Agostino-Pearson Omnibus Test

An alternative test of the same nature for samples larger than 50 was designed by D'Agostino (1973).

To assess the symmetry or asymmetry generally, skewness is measured and to evaluate the shape of the distribution kurtosis is overlooked. D'Agostino-Pearson (1973) test standing based on skewness and kurtosis test and these are also assessing through moments. The equation is given as:

$$K^2 = Z^2(\sqrt{b_1}) + Z^2(b_2) \quad (4)$$

where $Z^2(\sqrt{b_1})$ and $Z^2(b_2)$ are the normal approximation equivalent to $\sqrt{b_1}$ and b_2 are sample skewness and kurtosis respectively. This statistic follows a chi-squared distribution with two degrees of freedom if the population is from a normal distribution. A large value K^2 leads to the rejection of the normality assumption.

For the skewness test, the skewness coefficient of a time series $X(t)$ is estimated as follows (Karamouz *et al.*, 2003):

$$\hat{S} = \frac{\frac{1}{N} \sum_{i=1}^N (X_{(t)} - \bar{X})^3}{\left[\frac{1}{N} \sum_{i=1}^N (X_{(t)} - \bar{X})^2 \right]^{\frac{3}{2}}} \quad (5)$$

where N is the number of sample data and \bar{X} is the sample mean for time series $X(t)$. The skewness test is based on the fact that the skewness coefficient of a normal variable is zero. If the series is normally distributed, \hat{S} is asymptotically normally distributed with the mean of zero, variance of $6/N$, hence, $(1-\alpha) \times 100\%$ confidence limit on skewness is defined as,

$$S \in \left[-Z_{\left(\frac{\alpha}{2}\right)} \sqrt{\frac{6}{N}}, Z_{\left(\frac{\alpha}{2}\right)} \sqrt{\frac{6}{N}} \right] \quad (6)$$

where $Z_{(\alpha/2)}$ is the $(1-\alpha/2)$ quantile of the standard normal distribution. Therefore, if \hat{S} falls within the limits of Eq.6, the hypothesis of normality cannot be rejected. The test is found to be reasonably accurate for $N > 150$.

The Shapiro-Wilk (S-W) test, D'Agostino-Pearson omnibus test, and skewness values are also displayed as the output of descriptive statistic explained in the last section and the result is presented in Section 3.1.2.

2.4. Data analysis

2.4.1. Test for homogeneity

Homogeneity tests help in assessing trend reliability and identifying suitable sub-periods for the analysis. Homogeneity test was carried out by use of Pettit's test, Standard Normal Homogeneity Test (SNHT), Buishand's test, Von Neumann Ratio which was implemented by the XLSTAT software.

The basis of these tests corresponds to the alternative hypothesis of a single shift. For all tests, p-values were being calculated using Monte Carlo resampling. Test of hypothesis would guide in taking a decision on from results of homogeneity on the tables. The hypothesis is stated as follows:

Null hypothesis: H_0 : Data are Homogeneous

Alternative hypothesis: H_a There is a date at which there is a change in the data

Test interpretation:

H_0 : Data are homogeneous

H_a : There is a date at which there is a change in the data. The p-value shows that the null hypothesis is rejected; we can conclude that there is a shift between two parts of our time series. The associated pilot confirms this result. The SNHT test (Standard Normal Homogeneity Test) is usually applied to a series of ratios that compare the observations with an average. The ratios are then standardized. The null hypothesis are: H_0 : The obtained ratios follow an $N(0,1)$ distribution.

Since the p -value is very small, we reject the null hypothesis and thus conclude that there exists a shift in the time series. This result confirms the result of the first test. The Buishand's test can be used on variables following any type of distribution. It is based on the null hypothesis: H_0 : The T variables follow one or more distributions that have the same mean. Since the p -value is very small, this hypothesis is rejected and the alternative hypothesis will be: there exists a time t from which the variables change of mean. Finally, the von Neumann ratio is based on the sum of the square's differences between each pair of following time measures. The mean of this ratio is equal to 2 when the average of the time series is constant. The p -value is equal to 0.002, which leads us to reject the null hypothesis of homogeneity of the time series. This also confirms the preceding results. Von Neumann does not determine the time of change.

When p is smaller than the specified significance level, e.g. 0.05, the null hypothesis is rejected. In other words, if a significant change point exists, and the time series was divided into two parts at the location of the change point. For all these four tests, if the test statistic exceeds the critical value at a certain confidence level, the null hypothesis will be rejected at that confidence level.

The statistical analyses of every climatic time series must always be carried out for studying important time series characters, i.e., normality, homogeneity, seasonality, presence of trends and changes, etc.

2.4.2. Pettitt's test (Non-Parametric Rank Test)

This test developed by Pettitt is a nonparametric test, which is useful for evaluating the occurrence of abrupt changes in climatic records (Smadi and Zghoul, 2006). One of the reasons for using this test is that it is more sensitive to breaks in the middle of the time series (Wijngaard *et al.*, 2003). Pettitt's test is a nonparametric test that requires no assumption about the distribution of data. Pettitt's test is an adaptation of the rank-based Mann-Whitney test that allows identifying the time at which the shift occurs. The statistic used for Pettitt's test is computed as follows: The null and alternative hypotheses in this test are the same as in the Buishand test, and this test is also more sensitive to the breaks in the middle of the series (Costa and Soares, 2009). The ranks $r_1 \dots r_n$ of the $Y_1 \dots Y_n$ is used to calculate the statistics.

$$X_k = 2 \sum_{i=1}^k r_i - k(n+1) \quad k = 1, 2, \dots, n \quad (7)$$

If a break occurs in year K , then the statistic is maximal or minimal near the year $k = K$:

$$X_k = \max_{1 \leq k \leq n} |X_k| \quad (8)$$

2.4.3. Standard Normal Homogeneity Test (SNHT)

SNHT is one of the most popular homogeneity tests in climate studies. The null and alternative hypotheses in this test are the same as in the Buishand test; however, unlike the Buishand test, SNHT is more sensitive to the breaks near the beginning and the end of the series (Costa and Soares, 2009). Alexandersson and Moberg (1997) proposed a statistic $T(k)$ to compare the mean of the first k years of the record with that of the last $(n - k)$ years:

$$T(k) = kz_1^{-2} + (n - k)z_2^{-2} \quad k = 1, 2, \dots, n \quad (9)$$

$$\text{Where, } \bar{z}_1 = \frac{1}{n} \frac{\sum_{i=1}^k (Y_i - \bar{Y})}{S} \text{ and } \bar{z}_2 = \frac{1}{n - k} \frac{\sum_{i=k+1}^n (Y_i - \bar{Y})}{S} \text{ and } S = \frac{1}{n} \sum_{i=1}^k (Y_i - \bar{Y}) \quad (10)$$

If a break is located at the year K , then $T(k)$ reaches a maximum near the year $k = K$. The test statistic T_0 is defined as:

$$T_0 = \max_{1 \leq k \leq n} T(k) \quad (11)$$

The null hypothesis is rejected if T_0 is above a certain level, which is dependent on the sample size.

2.4.4. Buishand's test (Parametric test)

This test is used to detect a change in the mean by studying the cumulative deviation from the mean. It bases on the adjusted partial sums or cumulative deviation from the mean. According to Al-Ghazali, and Alawadi (2014) the null hypothesis is that the data are homogenous and normally distributed and the alternative hypothesis is that there is a date at which a change in a mean occurs.

This test supposes that tested values are independent and identically normally distributed (null hypothesis). The alternative hypothesis assumes that the series has a jump-like shift (break). This test is more sensitive to breaks in the middle of the time series (Costa and Soares, 2009). The test statistics, which are the adjusted partial sums (Buishand, 1982), are defined as $S_0^* = 0$ and

$$S_k^* = \frac{n \sum_{i=1}^k (Y_i - \bar{Y})}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad k = 1, 2, \dots, n \quad (12)$$

When series are homogeneous, the values of S_k^* will fluctuate around zero because no systematic deviations of the Y_i values concerning their mean will appear. Q-statistics: if a break is present in year K , then S_k^* reaches a maximum (negative shift) or minimum (positive shift) near the year $k = K$.

$$Q = \max_{0 \leq k \leq n} S_k^* \quad (13)$$

R-statistics (Range Statistics) are,

$$R = \left(\max_{0 \leq k \leq n} S_k^* - \min_{0 \leq k \leq n} S_k^* \right) \quad (14)$$

Buishand (1982) gives critical values for Q and R for different data set lengths random values; the alternative hypothesis is that the values in the series are not randomly distributed.

2.4.5. Von Neumann Ratio Test (Non-Parametric Test)

Von Neumann proposed a nonparametric test where the statistic is defined as the ratio of the mean square successive (year-to-year) difference to the variance. The null hypothesis is that the data are independent, identically distributed random quantities and the alternative is that the time series is not randomly distributed. Under the null hypothesis of a constant mean, the expected value of the test statistic is equal to two. The von Neumann ratio test is not location-specific, which means that it gives no information about the date of the break.

In this test, the null hypothesis is that the data are independent identically distributed. The von Neumann ratio N is defined as the ratio of the mean square successive (year to year) difference to the variance (von Neumann):

$$N = \frac{\sum_{i=1}^{n-1} (Y_i - Y_{i+1})^2}{\sum_{i=1}^n (Y_i - \bar{Y})^2} \quad (15)$$

Hereafter, for each of the test descriptions, n is the data set length, Y_i is the i^{th} element of the data set, \bar{Y} is the mean value of the data set. When the sample is homogeneous the expected value is $N = 2$. If the sample contains a break, then the value of N tends to be lower than this expected value. If the sample has rapid variations in the mean, then values of N may rise above two (Klein, 2007). This test gives no information about the location of the shift. Critical values for N for different data set lengths.

XLSTAT statistical software used a hypothesis testing method to detect homogeneity of the rainfall data. This software was used to execute the homogeneity analysis. The results were categorized into three classes, which are useful, doubtful, and suspect according to the number of tests rejecting the null hypothesis. Based on an alpha value of 0.05 (95% significance level), for p -value bigger than alpha value, the series was considered to be homogeneous.

3.0. Results and Discussion

3.1. Test result for reliability of CRU data (preliminary data validation)

The descriptive statistics of NiMeT and CRU mean annual rainfall and temperature are presented in Tables 4 and 5. Table 4 shows that the mean annual rainfall varies from 3074.098 mm in Bayelsa and 1263.575 mm in Lagos. The standard deviation varies from 96.9369mm to 160.7266mm while the skewness and Kurtosis vary from 0.06751 to -0.01484 and -1.61271 to -0.57684 respectively. These values of skewness and Kurtosis are indicative that the rainfall series approximate to normal distribution.

Table 4: Descriptive statistics of NiMeT and CRU mean annual rainfall (1956- 1986 and 1987-2016)

Climatic Station	Climatic Data	Mean mm	SD mm	Variance	Min Mm	Max mm	Range mm	Sum mm	Skewness	Kurtosis
Akwa Ibom	NIMET	377.4364	235.4844	55452.89	35	800.125	765.125	4151.8	0.201617	-0.78341
	CRU	236.5422	150.7492	22725.32	34.13	458.181	424.049	2838.5	-0.01484	-1.55418
Bayelsa	NIMET	237.1409	139.3913	19429.94	33.27	428.725	395.45	2608.55	-0.2039	-1.55741
	CRU	256.1748	160.7266	25833.04	45.39	464.289	418.903	3074.098	0.067378	-1.51304
Cross River	NIMET	257.3424	147.4242	21733.91	23	449.8	426.8	2830.767	-0.3454	-1.16572
	CRU	230.2533	153.8214	23661.02	22.58	448.109	425.52	2763.039	-0.0152	-1.61271
Delta	NIMET	176.484	120.079	14418.97	4.1	377.63	373.53	1941.33	-0.1477	-0.8333
	CRU	206.0484	142.8669	20410.94	25.08	404.26	379.176	2472.581	0.037337	-1.60228
Edo	NIMET	191.7341	128.3811	16481.72	0.55	413.375	412.825	2109.075	-0.04404	-0.74626
	CRU	170.348	117.3442	13769.66	16.55	335.84	319.286	2044.176	-0.09218	-1.60147
Lagos	NIMET	114.870	0.88266	5024.352	9.55	241.75	232.2	1263.575	0.443287	-0.58833
	CRU	136.2877	99.56	9912.54	15.364	331.068	315.70	1635.453	0.504882	-0.57684
Ogun	NIMET	114.7036	79.180	6269.618	2.86	209.74	206.88	12	-0.32077	-1.79549
	CRU	140.0217	96.9369	9396.777	13.88	280.636	266.76	1680.26	0.06751	-1.53347
Ondo	NIMET	119.3667	67.100	4502.435	0	221.5	221.5	1313.03	-0.50953	-0.46507
	CRU	167.3427	116.23	13509.37	17.12	344.131	327.013	2008.1	0.055487	-1.43844
Rivers	NIMET	215.2455	133.545	17834.28	20.35	344.1312	369.75	2367.7	0.039566	-1.66506
	CRU	246.76	154.04	23729.21	44.17	456.724	412.55	2961.13	0.030139	-1.58832

Table 5: Descriptive statistics of NiMeT and CTU mean annual temperature (1956- 1986 and 1987-2016)

Climatic Station	Climatic Data	Mean mm	SD mm	Variance	Min Mm	Max mm	Range mm	Sum mm	Skewness	Kurtosis
Akwa Ibom	NIMET	27.074	1.0575	1.118	25.775	28.891	3.116	297.8	0.482552	-0.8468
	CRU	30.569	1.328	1.764	28.31	32.332	4.022	366.83	-0.32776	-1.00893
Bayelsa	NIMET	26.66	0.9298	0.864	25.4	28	2.6	293.3	-0.03137	-1.3631
	CRU	30.39	1.3242	1.7534	28.267	32.046	3.7786	364.717	-0.43492	-1.31231
Cross River	NIMET	26.86	1.1483	1.3187	25.24	28.77	3.525	295.46	0.242124	-0.92303
	CRU	30.562	1.431	2.0476	28.064	32.490	4.43	366.74	-0.33311	-0.85852
Delta	NIMET	27.45	1.1794	1.391	25.687	29.025	3.338	301.98	-0.2685	-1.1837
	CRU	30.579	1.561	2.438	28.077	32.501	4.42	366.95	-0.45921	-1.31496
Edo	NIMET	27.27	1.4263	2.035	25.068	29.068	4	300.05	-0.31898	-1.39383
	CRU	30.447	1.838	3.378	27.507	32.729	5.22	365.366	-0.40085	-1.34195
Lagos	NIMET	26.96	1.1604	-1.3419	25.2	28.6	3.4	296.6	-0.11815	-1.42984
	CRU	30.597	1.629	2.6558	28.046	32.490	4.44	367.162	-0.43246	-1.4675
Ogun	NIMET	27.08	1.336	1.786	25.1	29.1	4	297.9	-0.02428	-1.2389
	CRU	30.81	1.856	3.444	27.826	33.063	5.24	369.73	-0.3822	-1.444
Ondo	NIMET	25.38	1.415	2.003	23.2	27.2	4	279.2	-0.26792	-1.38252
	CRU	30.664	1.8259	3.3339	27.746	32.949	5.20	367.97	-0.368	-1.391
Rivers	NIMET	27.017	0.972976	0.947	25.76	28.508	2.75	297.187	0.2653	-1.2259
	CRU	30.512	1.299	1.688	28.368	32.157	3.79	366.14	-0.3456	-1.1793

The descriptive statistics of mean annual temperature (1956-2016), in the coastal region of Nigeria are presented in Table 5. The table was obtained by use of XL Statistic software as explained in Section 3.3.1. The table shows that the mean annual temperature varies from 369.73°C in Ogun and 293.3°C in Bayelsa. The standard deviation varied from 1.856°C to 1.3242°C while the skewness and Kurtosis varied from -0.45921 to -0.32776 and -1.4675 to -0.85852 respectively. These values of skewness and Kurtosis are indicative that the rainfall series approximate to normal distribution.

3.2. Result for differences in mean annual rainfall during the two climatic periods (simple t-Test)

The results of values obtained from the computer output are presented in Table 6. The table depicts the values of the means in the two climatic periods in the states for which the Student t-test was used to establish if there were differences between them. The Pearson correlation, the number of observations (n) the degree of freedom ($df = n-1$) are shown. The computed t statistics, as well as the t critical for one tail and two-tail test, are also presented.

The hypothesis is stated as follows:

The null hypothesis is that there is no statistical difference in mean seasonal rainfall distribution between the NIMET and CRU data. The alternate hypothesis is that there is a statistical difference in mean seasonal rainfall distribution between the NIMET and CRU data:

Null hypothesis $H_0: \mu_1 = \mu_2$

While the alternative hypothesis is $H_a: \mu_1 \neq \mu_2$ i.e. $\mu_1 > \mu_2$ or $\mu_1 < \mu_2$.

Hence, two-tail tests are used for the analysis. Based on the t statistic test statistic and the degrees of freedom, we determine the P -value. The P -value is the probability that a t statistic having 11 degrees of freedom is more extreme than 1.725. Since this is a two-tailed test, "more extreme" means greater than 1.725 or less than -1.725 i.e. $-1.725 < T_0 < 1.725$.

Table 6: Results of Simple t - test on comparison of NIMET and CRU data for rainfall

Location	Simple t-test (T_0) Computed	Df (n-2)	Alpha	P-value (Two tail)	Decision
AKWA IBOM	1.660	22	0.05	1.725	Cannot Reject H_0
BAYELSA	0.295	22	0.05	1.725	Cannot Reject H_0
CROSS RIVER	0.420	22	0.05	1.725	Cannot Reject H_0
DELTA	0.523	22	0.05	1.725	Cannot Reject H_0
EDO	0.406	22	0.05	1.725	Cannot Reject H_0
LAGOS	0.429	22	0.05	1.725	Cannot Reject H_0
OGUN	0.569	22	0.05	1.725	Cannot Reject H_0
ONDO	1.181	22	0.05	1.725	Cannot Reject H_0
RIVERS	0.511	22	0.05	1.725	Cannot Reject H_0

Hence, there is no significant statistical difference in mean of seasonal rainfall distribution between the NIMET and CRU data at a significance level of 0.05.

To determine the suitability/ adequacy of the CRU rainfall and temperature data for sixty-one years (1956-2017) for further analysis, the data were further subjected to validation with observed NIMET data covering the same period by comparison of their descriptive statistics using simple t test and the goodness of fit criterion such as Coefficient of Determination (R^2), adjusted- R^2 , Mallows' Process Capability (C_p), Kendall Tau (K tau), Root Mean Squared (RSR) and Mean Absolute Percentage Error (MAPE). This is shown in Table 7 and 8 for rainfall and temperature respectively.

Table 7: Comparison between NIMET and CRU data for rainfall

State	R^2	Rating	Adj R^2	Rating	K tau	Rating	RSR	Rating	MAPE	Rating	C_p	Rating
Akwa Ibom	0.924	V.good	0.901	V.good	0.782	V.good	0.315	V.good	18.6	Good	2	OK
Bayelsa	0.931	V.good	0.907	V.good	0.782	V.good	0.302	V.good	19.2	Good	2.8	OK
Cross River	0.935	V.good	0.912	V.good	0.801	V.good	0.293	V.good	30.3	R	2	OK
Delta	0.865	V.good	0.83	V.good	0.818	V.good	0.41	V.good	66	I	2	OK
Edo	0.890	V.good	0.867	V.good	0.709	V.good	0.369	V.good	23.48	R	2	OK
Lagos	0.774	V.good	0.709	V.good	0.709	V.good	0.539	Good	44.79	R	2	OK
Ogun	0.888	V.good	0.850	V.good	0.745	V.good	0.381	V.good	50	R	2	OK
Ondo	0.882	V.good	0.869	V.good	0.782	V.good	0.362	V.good	15.37	Good	2	OK
River	0.927	V.good	0.903	V.good	0.818	V.good	0.312	V.good	18.17	Good	2	OK

R = Reasonable, I = Inaccurate, OK = capable or Reliable

Table 8: Comparison between NIMET and CRU data for temperature

State	R^2	Rating	Adj R^2	Rating	K tau	Rating	RSR	Rating	MAPE	Rating	C_p	Rating
Akwa Ibom	0.933	V.good	0.925	V.good	0.927	V.good	0.272	V.good	0.679	H	2	OK
Bayelsa	0.954	V.good	0.949	V.good	0.891	V.good	0.225	V.good	0.588	H	2	OK
Cross River	0.973	V.good	0.97	V.good	0.964	V.good	0.173	V.good	0.523	H	2	OK
Delta	0.949	V.good	0.94	V.good	0.891	V.good	0.227	V.good	0.659	H	2	OK
Edo	0.976	V.good	0.973	V.good	0.855	V.good	0.165	V.good	0.634	H	2	OK
Lagos	0.931	V.good	0.923	V.good	0.823	V.good	0.227	V.good	0.987	H	2	OK
Ogun	0.946	V.good	0.94	V.good	0.844	V.good	0.244	V.good	0.772	H	2	OK
Ondo	0.883	V.good	0.87	V.good	0.771	V.good	0.36	V.good	1.379	H	2	OK
River	0.905	V.good	0.922	V.good	0.855	V.good	0.278	V.good	0.623	H	2	OK

H = Highly Reliable, OK = capable or Reliable

Tables 7 and 8 shows the values of R^2 obtained from the comparison. They range from 0.935 (Cross River) to 0.774 (Lagos) for rainfall and from 0.976 (Edo) to 0.883 (Ondo) for temperature. Following Table 2, the performance rating of R^2 is very good as $0.75 < R^2 \leq 1$. With the minimum value of 0.77788 for both rainfall and temperature, the performance rating of R^2 in validating the CRU data is very good.

The results of this validity analysis indicate that the CRU data obtained is very reliable and thus justified their use for the further analysis carried out in the study.

RMSE close to zero and R-Square value approaching 1 is indicative of high accuracy between observed and predicted values. Based on a rule of thumb, it can be said that RMSE values between 0.2 and 0.5 show that the model can relatively predict the data accurately.

C_p , processes capability sometimes referred to as *Mallows' C_p* , shows whether the distribution can potentially fit inside the specification. C_p is an index used to assess the width of the process spread in comparison to the width of the specification. It is calculated by dividing the allowable spread by the actual spread. The allowable spread is the difference between the upper and lower specification limits. The actual spread is 6 times the estimated standard deviation.

$$C_p = (USL - LSL)/6\sigma \quad (16)$$

Where USL and LSL are upper and lower specification limits, σ estimated standard deviation.

3.3. Normality tests

Three common normality tests were carried out, namely the Shapiro Wilk Test (SWT), D'Agostino-Pearson Test, and Skewness Test. The results are presented in Tables 9 to 12. When the p -value is larger than the significance level, the decision is to fail to reject the null hypothesis because we do not have enough evidence to conclude that our data do not follow a normal distribution. The results indicate that the three normality tests were in agreement that rainfall series follow a normal distribution. Normal distribution assumption is very crucial for the reliability of results especially for parametric tests. The tables show values of test statistics, the p -values, and significant level $\alpha = 0.05$. It also provided for confirmation of normality as it indicated YES. To determine whether the data do not follow a normal distribution, we compare the p -value to the significance level. Usually, a significance level (denoted as α or alpha) of 0.05 works well. A significance level of 0.05 indicates that the risk of concluding the data do not follow a normal distribution—when the data do follow a normal distribution—is 5%.

Table 9: Results of test for normality of spatial rainfall data (normality acceptable if $-0.59 < s < 0.59$)

Location	SHAPIRO WILK TEST (SWT)				D'AGOSTINO-PEARSON TEST				SKEWNESS TEST		
	W STAT	P value	Alpha	normal	DA STAT	P value	alpha	normal	S	alpha	YES
Akwa Ibom	0.93868	0.50511	0.05	YES	02.0182	0.36455	0.05	YES	-0.01484	0.05	YES
Bayelsa	0.91194	0.25706	0.05	YES	1.9496	0.3773	0.05	YES	0.067378	0.05	YES
Cross River	0.92222	0.33758	0.05	YES	2.438	0.2955	0.05	YES	-0.0152	0.05	YES
Delta	0.92559	0.36799	0.05	YES	2.2248	0.3287	0.05	YES	0.037337	0.05	YES
Edo	0.92490	0.36169	0.05	YES	1.985	0.3706	0.05	YES	-0.09218	0.05	YES
Lagos	0.96089	0.7828	0.05	YES	0.5076	0.7758	0.05	YES	0.504882	0.05	yes
Ogun	0.94230	0.54799	0.05	YES	1.9859	0.3704	0.05	YES	0.06751	0.05	YES
Ondo	0.94946	0.63700	0.05	YES	1.4617	0.4815	0.05	YES	0.055487	0.05	YES
Rivers	0.92684	0.37975	0.05	YES	2.2976	0.3170	0.05	YES	0.030139	0.05	YES

Table 10: Results of test for normality of spatial temperature data (normality acceptable if $-0.59 < s < 0.59$)

Location	SHAPIRO WILK TEST (SWT)				D'AGOSTINO-PEARSON TEST				SKEWNESS TEST		
	W STAT	P-value	Alpha (α)	Normal ($p > \alpha$)	DA STAT	P value	alpha	Normal ($p > \alpha$)	S	alpha	YES
Akwa Ibom	0.94954	0.66381	0.05	YES	1.2855	0.5258	0.05	YES	-0.32776	0.05	YES
Bayelsa	0.91966	0.3158	0.05	YES	2.5974	0.2729	0.05	YES	-0.43492	0.05	YES
Cross River	0.95524	0.71132	0.05	YES	0.9932	0.6086	0.05	YES	-0.33311	0.05	YES
Delta	0.9069	0.2242	0.05	YES	2.7698	0.2503	0.05	YES	-0.45921	0.05	YES
Edo	0.9123	0.2594	0.05	YES	2.756	0.2520	0.05	YES	-0.40085	0.05	YES
Lagos	0.9080	0.23108	0.05	YES	3.009	0.2221	0.05	YES	-0.43246	0.05	YES
Ogun	0.9136	0.2689	0.05	YES	2.899	0.2346	0.05	YES	-0.38228	0.05	YES
Ondo	0.9143	0.274	0.05	YES	2.864	0.2387	0.05	YES	-0.36835	0.05	YES
Rivers	0.9593	0.7775	0.05	YES	1.094	0.5787	0.05	YES	-0.34561	0.05	YES

The five tests scored H_0 for Edo, Lagos, Ogun, and Ondo state Rainfall data and H_a for Cross Rivers and Rivers States rainfall data. Three tests scored H_a and two tests scored H_0 for Akwa Ibom rainfall data. The average score could be considered as H_a . Also, three tests scored H_0 and two tests scored H_a for Bayelsa and Delta States rainfall data. The average score could be considered as H_0 .

Table 11a: Summary of homogeneity test for rainfall data (Pettit's test and SNHT)

State	PETTIT'S TEST				SNHT			
	K-Value	Year	P-Value	T	To-Value	Year	P-Value	T
Akwa Ibom	321	1971	0.17	Ho	9.894	1969	0.020	Ha
Bayelsa	243	1969	0.62	Ho	9.39	2011	0.032	Ha
Cross River	618	198	0.00	Ha	21.44	1984	0.000	Ha
Delta	258	1980	0.48	Ho	8.917	2012	0.038	Ha
Edo	149	2011	0.32	Ho	3.429	2011	0.618	Ho
Lagos	278	1970	0.36	Ho	7.883	1970	0.157	Ho
Ogun	192	1986	0.83	Ho	4.131	1970	0.414	Ho
Ondo	244	1986	0.54	Ho	3.033	1958	0.701	Ho
Rivers	466	1980	0.00	Ha	14.12	1969	0.002	Ha

Table 11b: Summary of homogeneity test for rainfall data (Bush and Von Neumann test)

State	BUISHAND TEST							VON NEUMANN		
	Q-Value	Year	P-Value	T	R-Value	P-Value	T	N	P-Value	T
Akwa Ibom	10.303	1971	0.04	Ha	11.94	0.065	Ho	1.26	0.002	Ha
Bayelsa	6.616	2011	0.369	Ho	8.227	0.547	Ho	1.42		Ha
Cross River	18.04	1984	0.000	Ha	18.043	0.000	Ha	0.000	0.00	Ha
Delta	7.669	1999	0.216	Ho	9.617	0.238	Ho	1.284	0.002	Ha
Edo	3.998	20.11	0.895	Ho	7.572	0.669	Ho	1.675	0.10	Ho
Lagos	9.276	1970	0.062	Ho	11.842	0.028	Ha	1.679	0.088	Ho
Ogun	6.715	1970	0.346	Ho	10.151	0.193	Ho	1.618	0.061	Ho
Ondo	5.835	1986	0.509	Ho	8.625	0.452	Ho	1.655	0.098	Ho
Rivers	13.33	1980	0.002	Ha	13.899	0.011	Ha	1.124	0.00	Ha

Table 12a: Summary of homogeneity test for temperature data

Location	PETTIT'S TEST			SNHT		
	K-Value	Year	Trend	To-Value	Year	Trend
Akwa Ibom	652.00	1980	Ha	23.011	1980	Ha
Bayelsa	708.00	1980	Ha	27.490	1980	Ha
Cross River	648.00	1986	Ha	22.561	1980	Ha
Delta	718.00	1980	Ha	28.073	1980	Ha
Edo	708.00	1980	Ha	27.287	1980	Ha
Lagos	630.00	1980	Ha	20.504	1980	Ha
Ogun	642.00	1980	Ha	21.647	1980	Ha
Ondo	684.00	1980	Ha	24.723	1980	Ha
Rivers	688.00	1980	Ha	25.900	1980	Ha

Table 12b: Summary of homogeneity test for temperature data

Location	BUISHAND TEST			VON NEUMANN			
	Q-Value	Year	Trend	R-Value	Trend	N	
Akwa Ibom	18.357	1980	Ha	18.357	Ha	1.078	Ha
Bayelsa	20.064	1980	Ha	20.064	Ha	0.995	Ha
Cross River	18.177	1980	Ha	18.177	Ha	1.100	Ha
Delta	20.278	1980	Ha	20.276	Ha	0.992	Ha
Edo	19.990	1980	Ha	19.990	Ha	1.001	Ha
Lagos	17.328	1980	Ha	17.418	Ha	1.134	Ha
Ogun	17.805	1980	Ha	17.886	Ha	1.096	Ha
Ondo	19.027	1980	Ha	19.027	Ha	1.016	Ha
Rivers	19.475	1980	Ha	19.475	Ha	0.978	Ha

Test interpretations of rainfall data continued

Case 1	Case 2
H_0 : Data are Homogeneous H_a : There is a date at which there is a change in the data. As the computed p -value is greater than the significance level $\alpha = 0.05$, one cannot reject the null hypothesis H_0 . We have to accept the null hypothesis that the data is homogeneous	H_0 : Data are Homogeneous H_a : There is a date at which there is a change in the data. As the computed p -value is lower than the significance level $\alpha = 0.05$, one should reject the null hypothesis H_0 , and accept the alternate hypothesis H_a that the data is inhomogeneous

3.3.1. Test interpretations of temperature data

The five tests scored H_a for all the States in the study area for temperature data

H_0 : Data are Homogeneous

H_a : There is a date at which there is a change in the data

As the computed p -value (<0.0001) is lower than the significance level $\alpha = 0.05$, one should reject the null hypothesis H_0 , and accept the alternate hypothesis H_a that the data is not homogeneous.

CRU TS is not specifically homogeneous. Some National Meteorological Agencies (NMAs) homogenize their station observations, either before release or at a later stage (requiring a re-release). Therefore, many CRU TS observations have been homogenized (and also quality controlled) within each country. However, performing additional homogenization on the CRU TS databases would be complicated and not completely possible because of elements of the process, such as partly synthetic variables and the use of published climatology. The result obtained from homogeneity test is presented in Figures 2a to 2j and Table 13.

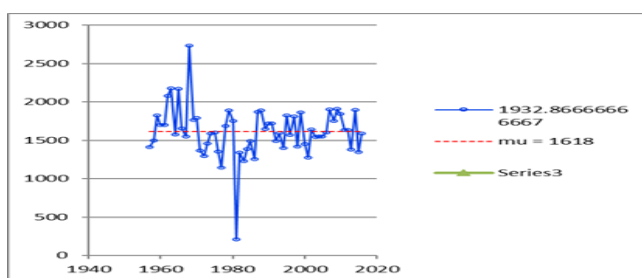


Figure 2a: Lagos State rainfall Pettitt test

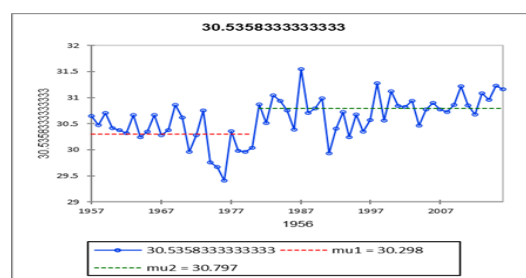


Figure 2b: Lagos State temperature Pettitt Test

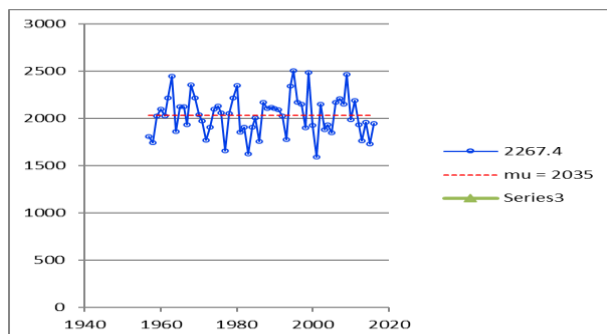


Figure 2c: Edo State rainfall Pettitt test

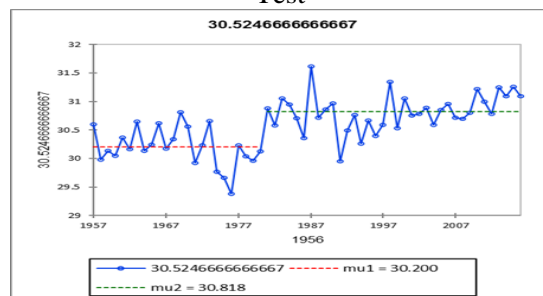


Figure 2d: Edo State temperature Pettitt test

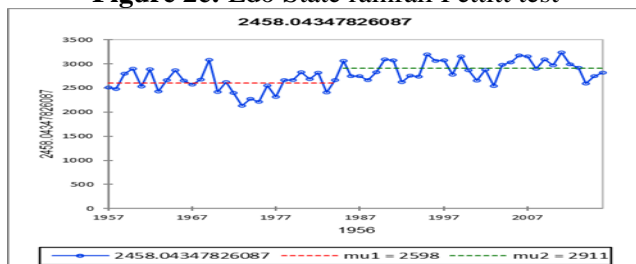


Figure 2e: Cross River State rainfall Pettitt test

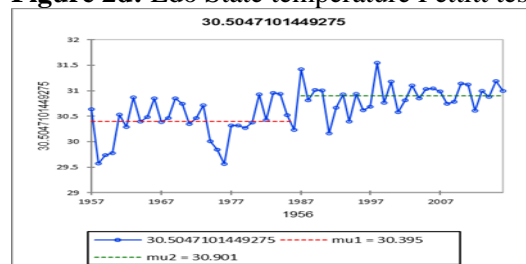


Figure 2f: Cross River State temperature Pettitt test

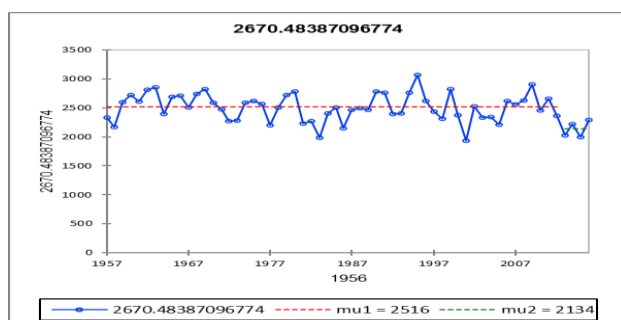


Figure 2g: Delta State rainfall SNHT test

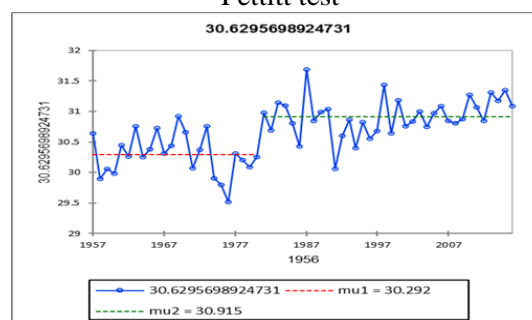


Figure 2h: Delta State temperature SNHT test

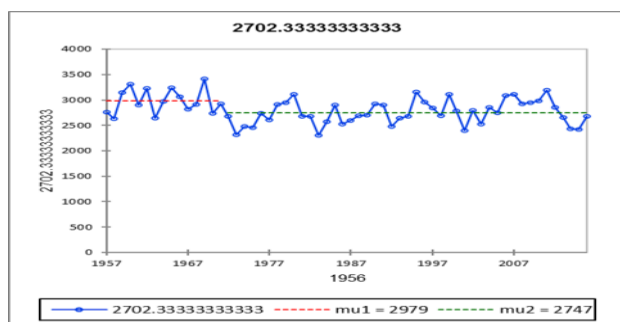


Figure 2i: Akwa Ibom State rainfall Buishand's test

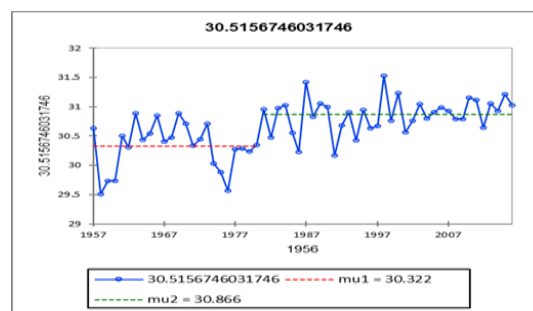


Figure 2j: Akwa Ibom State temperature Buishand's test

Table 13: Summary of abrupt change result obtained from homogeneity test (Figure 2a to 2j)

State	Year of abrupt change for rainfall	Year of abrupt change for temperature
Akwa Ibom	1971	1980
Bayelsa	2011	1980
Cross River	1984	1986
Delta	1980	1980
Edo	2011	1980
Lagos	1970	1980
Ondo	1970	1980
Ogun	1986	1980
Rivers	1980	1980

4.0. Conclusion

The five homogeneity tests conducted indicated that CRU rainfall data for Bayelsa, Delta Edo, Lagos, Ogun, and Ondo states are homogeneous those for Akwa Ibom, Cross Rivers, and Rivers States are inhomogeneous. Also, the five homogeneity tests conducted indicated that CRU temperature data for all the States in the study area are inhomogeneous.

National Meteorological Agencies (NMAs) homogenize their station observations, either before release or at a later stage (requiring a re-release). Therefore, many CRU TS observations have been homogenized (and also quality controlled) within each country. Hence, inhomogeneity observed may not be due to data quality but to climate variability and climate change. For temperature, the year of significant abrupt change was likely in 1980 where there were breakpoints. The study also indicated that 1980 was the driest year in Lagos state.

References

- Agha, O. M. A., Bağcı, Ç. S. and Şarlak, N. (2017). Homogeneity analysis of precipitation series in North Iraq. *IOSR Journal of Applied Geology and Geophysics*, 5, pp. 57–63.
- Alexandersson, H. and Moberg, A. (1997). Homogenization of Swedish Temperature Data Part I: Homogeneity Test for Linear Trends. *International Journal of Climatology*, 17, pp. 25–34.
- Al-Ghazali, N. O. S. and Alawadi, D. A. H. (2014). Testing the Homogeneity Of Rainfall Records for Some Stations in Iraq. *International Journal of Civil Engineering and Technology (IJCIET)*, 5, pp. 76–87.
- Buishand, T. A. (1982). Some Methods for Testing the Homogeneity of Rainfall Records. *Journal of Hydrology*, 58, pp. 11–27.
- Costa, A. C. and Soares, A. (2009). Homogenization of climate data: review and new perspectives using geostatistics. *Math. Geosci.* 41, pp. 291–305.
- D' Agostino, R. and Pearson, E. S. (1973). Test for departure from normality: Empirical results for the distributions of b2 and -4131. *Biometrika*, 60, pp. 613–622.

Karamouz, M., Szidarovszky, F. and Zahraie, B. (2003). *Water Resources Systems Analysis*, Lewis Publishers, pp. 159-160.

Klein, T. A. M., Zwiers, F. W. and Zhang, X. (2009). *Guidelines on Analysis Adaptation*. Geneva: WMO.

Kidd, C., Becker, A., Huffman, G. J., Muller, C. L., Joe, P., Skofronick-Jackson, G. and Kirschbaum, D. B. (2017). So, how much of the Earth's surface is covered by rain gauges? *Bulletin of the American Meteorological Society*, 98, pp. 69–78.

Kuruk, P., (2004): Customary Water Laws and Practices: Nigeria. Available at: <http://www.fao.org/legal/advserv/FAOIUCNcs/Nigeria.pdf>

Ma, Z., Kang, S., Zhang, L., Tong, L. and Su, X. (2008). Analysis of impacts of climate variability and human activity on streamflow for a river basin in the arid region of northwest China. *Journal of Hydrology*, 352, 239-249.

Martínez, M. D., Serra, C., Burgueño, A. and Lana, X. (2009). Time Trends of Daily Maximum and Minimum Temperatures in Catalonia (Ne Spain) for the Period 1975–2004. *International Journal of Climatology*, 30(2), pp. 267–290.

Mitchell, T. D. and Jones, P. D. (2005). An Improved Method of Constructing A Database Of Monthly Climate Observations And Associated High-Resolution Grids. *International Journal Of Climatology Int. J. Climatol.* 25, pp. 693–712.

Nnamchi, H. C., Anyadike, R. N. C. and Emeribe, C. N. (2008). Spatial patterns of the Twentieth Century mean seasonal precipitation over West Africa. *Nigeria Journal of Space Research*, 6, pp 89-101.

Nwilo, P. C. and Badejo, O. T. (2006). Impacts and Management of Oil Spill Pollution along with the Nigerian Coastal Areas. *Administering Marine Spaces: International Issues*, 119, pp. 1-15.

Sarojini, B. B., Stott, P. A. and Black, E. (2016). Detection and attribution of human influence on regional precipitation. *Nat. Clim. Change*, 6(7), pp. 669–675.

Shapiro, S. S. and Wilk, M. B. (1965). An analysis of variance test for normality (complete samples). *Biometrika*, 52, pp. 591–611.

Smadi, M. M. and Zghoul, A. (2006). A Sudden Change in Rainfall characteristics in Amman, Jordan during the Mid 1950s. *Am J Environ Sci* 2(3), pp. 84–91.

Wijngaard, J. B., Klein Tank, A. M. G. and Können, G. P. (2003). Homogeneity of 20th Century European Daily Temperature and Precipitation Series 23, pp. 679–692.

Yozgatligil, C., Purutcuoglu, V., Yazici, C. and Batmaz, I. (2015). The validity of homogeneity tests for meteorological time series data: a simulation study, In *Proceedings of the 58th World Statistics Congress (ISI2011)*, Dublin, Ireland. . <http://2011.isiproceedings.org/papers/950636.pdf> (accessed 27 March 2015).

Cite this article as:

Agbonaye A. I. and Izinyon O. C.. 2021. Evaluating the Quality of Spatial Data for the Analysis of Climate Variability in the Coastal Region of Nigeria. *Nigerian Journal of Environmental Sciences and Technology*, 5(1), pp. 76-90. <https://doi.org/10.36263/nijest.2021.01.0236>